The Nature of Generic Cosmological Singularities[*]

Claes Uggla[†]

Department of Physics, University of Karlstad,
S-651 88 Karlstad, Sweden

Abstract

The existence of a singularity by definition implies a preferred scale—the affine parameter distance from/to the singularity of a causal geodesic that is used to define it. However, this variable scale is also captured by the expansion along the geodesic, and this can be used to obtain a regularized state space picture by means of a conformal transformation that factors out the expansion. This leads to the conformal 'Hubble-normalized' orthonormal frame approach which allows one to translate methods and results concerning spatially homogeneous models into the generic inhomogeneous context, which in turn enables one to derive the dynamical nature of generic cosmological singularities. Here we describe this approach and outline the derivation of the 'cosmological billiard attractor,' which describes the generic dynamical asymptotic behavior towards a generic spacelike singularity. We also compare the 'dynamical systems picture' resulting from this approach with other work on generic spacelike singularities: the metric approach of Belinskii, Lifschitz, and Khalatnikov, and the recent Iwasawa based Hamiltonian method used by Damour, Henneaux, and Nicolai; in particular we show that the cosmological billiards obtained by the latter and the cosmological billiard attractor form complementary 'dual' descriptions of the generic asymptotic dynamics of generic spacelike singularities.

[†]Electronic address: `Claes.Uggla@kau.se`

1 Introduction

The singularity theorems tell us that singularities occur under very general circumstances in general relativity (GR), but they say little about their nature—detailed insights require further analysis of Einstein's field equations. Of particular interest is the structure of generic singularities; it is probably fair to say that there have historically been three 'dominant' approaches to address this issue: (i) the metric approach used by Belinskii, Lifshitz, and Khalatnikov [1, 2, 3] (henceforth these authors, and their work, are referred to as BKL), (ii) the Hamiltonian approach, and (iii) the dynamical systems approach, which will be the main topic of the present paper, but let us begin with a brief historical review.

During the sixties, seventies, and early eighties, BKL performed a heuristic analysis that eventually resulted in the conjecture that a generic singularity is *spacelike*, *local*,[1] *vacuum dominated*, and *oscillatory*, when the matter source is a perfect fluid with a radiation equation of state, or more precisely: the time evolution of a solution in the vicinity of a generic singularity is described by a sequence of generalized Kasner solutions,[2] where one generalized Kasner solution is linked to the next by a generalized vacuum Bianchi type II solution, leading to an oscillatory behavior, discretely described by a chaotic Kasner map [4, 5, 6]. To reach these conclusions BKL (i) employed synchronous coordinates, i.e. Gaussian normal coordinates, such that the singularity occurred simultaneously, (ii) considered certain spatially homogeneous (SH) metrics, (iii) replaced constants with spatial functions, (iv) inserted the resulting expressions into Einstein's field equations as a starting point for a perturbative expansion which yielded a consistency check, which, after some difficulties, led to the above scenario.

In the late sixties and early seventies the BKL picture obtained heuristic support from results obtained by the Hamiltonian approach to asymptotic Bianchi type IX dynamics developed by Misner and Chitré [7, 8, 9]. In one variation of this approach the dynamical evolution towards the singularity is described in terms of free motion in an abstract flat Lorentzian 'minisuperspace' [3] surrounded by potential walls, or alternatively by means of a spatial projection in minisuper space which leads to free motion inside a potential well described by moving walls; for further use and developments of this picture, see the review by Jantzen [10] and [11] [Chapter 10]. In another version of the Hamiltonian approach Misner and Chitré [8, 9] described the asymptotic dynamics by projecting the motion in minisuperspace onto a region in hyperbolic space bounded by sharp walls, forming a 'billiard table', see [8, 9]. The Hamiltonian approach was later generalized so that the general inhomogeneous case could also be studied, culminating in the recent work by Damour, Henneaux, and Nicolai (aimed primarily at a broader string theoretic context), see [12, 13] and references therein.

Despite the indisputable ingenuity and power of the BKL and the Hamiltonian approaches, they have unfortunately yielded rather vague statements from a rigorous mathematical point of view, and this has led to considerable debate. But there exist exceptions in non-oscillatory cases: rigorous results have been obtained by Moncrief, Isenberg, Berger, and collaborators, who also obtained numerical support for the general basic BKL picture, see [14], and [15] for a review and additional references; note also the successful use of so-called Fuchsian methods as regards theorems about asymptotic non-oscillatory dynamics, particularly the result by Andersson and Rendall [16] for the general case of a stiff fluid or massless scalar field.

The last decade has seen a remarkable theoretical development in SH cosmology, largely in connection with the book "Dynamical Systems in Cosmology" [11]. The dynamics of e.g. perfect fluid models are now to a large extent understood in terms of rigorous statements, of which many have been elevated to mathematical theorems. Notably, Ringström obtained the first theorems about oscillatory behavior for Bianchi type VIII and, more substantially, type IX [17, 18] models. The

---

[1]Local means that near the singularity the spatial points essentially 'decouple,' in the sense that the dynamical evolution is asymptotically governed by ordinary differential equations with respect to time.

[2]Contrary to what the name implies, the generalized Kasner solution is not a solution to Einstein's equations; instead it is the line element one obtains if one replaces the constants in the vacuum Bianchi type I solution with spatially dependent functions.

[3]Superspace is the space of spatial metrics; minisuperspace is superspace restricted to the SH case.

progress in this area is due to the use of dynamical systems techniques applied to the 'dynamical systems approach' to SH cosmology, see [11, 19].

In an attempt to broaden the success to the general inhomogeneous case, Uggla, van Elst, Wainwright, and Ellis [20] introduced a dynamical systems formulation for Einstein's field equations without any symmetries. This led to a description of generic asymptotic dynamics towards a generic spacelike singularity in terms of an attractor, which resulted in mathematically precise conjectures. Furthermore, this formulation served as a basis for numerical investigations of generic singularities, which yielded additional support for the expected generic picture as well as the discovery of new phenomena and subsequent refinements [21, 22, 23, 24, 25]. Recently this work was given a more sound geometric foundation [26], and further steps were taken by Heinzle, Uggla, and Röhr [27] to sharpen and substantiate exact rigorous mathematical statements about generic asymptotic dynamics towards a generic spatial singularity—a development we focus on below.

2   The conformal Hubble-normalized dynamical systems approach

Let us consider a generic spacelike singularity, for simplicity located in the past, which motivates referring to it as a 'cosmological singularity.' In this 'cosmological' context it is natural to focus on temporal aspects and investigate spacetime changes along a timelike reference congruence that originates from the singularity. Let us assume that we are sufficiently close to the singularity so that the expansion $\theta = \boldsymbol{\nabla}_a u^a$ is positive, where u^a is the future-directed unit tangent vector field of the reference congruence and where $\boldsymbol{\nabla}_a$ is the covariant derivative associated with the physical metric $\mathbf{g}$.

Although the singularity theorems say little about the nature of singularities, the very definition of a singularity and the one dynamical input that goes into the theorems—the Raychaudhuri equation for the expansion[4]—provide clues to how one might continue in the quest to understand what happens in the vicinity of a generic spacelike singularity. The existence of a singularity by definition implies a prominent variable scale—the affine parameter distance from/to the singularity of a causal inextendible geodesic that is used to define it, or, alternatively, the expansion (cf. FRW cosmology where the Hubble variable $H = \frac{1}{3}\theta$ is frequently used to obtain a characteristic time scale, H^{-1}, from the initial singularity). However, the key role played by the Raychaudhuri equation in the singularity theorems suggests that the expansion is particularly important—especially so in the case of a generic spacelike singularity where the expansion is expected to blow up, so that Einstein's field equations break down.

Causal properties constitute an important feature in GR—particularly in the derivation of the singularity theorems—and there are good reasons to believe that asymptotic causal properties are absolutely crucial in the case of a generic spacelike singularity: We expect a generic spacelike singularity to be a scalar curvature singularity,[5] associated with increasing ultra strong gravity that focuses light in all directions as the singularity is approached, leading to *asymptotic silence*, which we define as the formation of particle horizons that shrink to zero size in all directions along any time line that approaches the singularity, thus increasingly prohibiting communication;[6] for various special examples of asymptotic silence and asymptotic silence-breaking, see Lim et al. [25].

In relativity, due to the causal structure that links space and time, there only exists a single dimensional unit, leaving us with one dimensional scale. The above discussion suggests that one should adapt to the asymptotic 'dominant' scale and causal properties and asymptotically *regularize* Einstein's equations by factoring out the expansion or, equivalently, the Hubble variable $H = \frac{1}{3}\theta$, so that one obtains dimensionless state space variables that take *finite* values and capture the 'essential' dynamics: the geometric and natural way to do this is by means of a conformal

---

[4]In the case of timelike geodesics; in the null geodesic case an analogous equation plays a similar role.

[5]A non-scalar curvature singularity requires fine tuning.

[6]One may also refer to this as an anti-Newtonian limit, since, loosely speaking, the light cones are flattened out onto spacelike surfaces in the Newtonian limit, while they here 'collapse' onto time lines.

transformation; since $\mathbf{g}$ and H^{-2} have dimensional weight ['length']^2 (or equivalently ['time']^2), the appropriate transformation is given by

$$\mathbf{g} = H^{-2}\,\mathbf{G}\,, \tag{1}$$

where $\mathbf{G}$ is a dimensionless unphysical metric.

To obtain a formulation that reduces to the Hubble-normalized dynamical systems approach in SH cosmology, used in e.g. [11], as a special case, we introduce: a conformal 'Hubble-normalized' orthonormal frame,

$$\mathbf{g} = H^{-2}\,\mathbf{G} = H^{-2}\,\eta_{ab}\,\mathbf{\Omega}^a\,\mathbf{\Omega}^b\,, \tag{2}$$

where $\eta_{ab} = \mathrm{diag}(-1,1,1,1)$, and $a,b = 0,1,2,3$; conformal orthonormal vector fields $\boldsymbol{\partial}_a$ dual to $\mathbf{\Omega}^a$, i.e., $\langle \mathbf{\Omega}^a, \boldsymbol{\partial}_b \rangle = \delta^a{}_b$; a non-rotating timelike reference congruence—asymptotically conformally geodesic along a generic time line[7]—to which we adapt the frame, i.e., $\boldsymbol{\partial}_0 = H^{-1}\mathbf{u}$ is aligned tangentially to the time lines of the reference congruence, and the shift vector is therefore set to zero.

This then yields

$$\boldsymbol{\partial}_0 = \mathcal{N}^{-1}\,\partial_{x^0} = \mathcal{N}^{-1}\,\partial_0\,, \qquad \boldsymbol{\partial}_\alpha = E_\alpha{}^i \partial_{x^i} = E_\alpha{}^i \partial_i\,, \qquad \alpha = 1,2,3\,; \quad i = 1,2,3\,, \tag{3}$$

where $\mathcal{N}$ is the conformal lapse and $E_\alpha{}^i$ are the conformal spatial frame vector components, related to corresponding objects of $\mathbf{g}$ by

$$\mathcal{N} = H\,N\,, \qquad E_\alpha{}^i = e_\alpha{}^i/H\,. \tag{4}$$

Note that the components $E_\alpha{}^i$ are associated with the *contravariant* spatial 3-metric of $\mathbf{G}$: $G^{ij} = \delta^{\alpha\beta} E_\alpha{}^i E_\beta{}^j$; for further details, see [26].

The next ingredient in the 'conformal Hubble-normalized orthonormal frame approach' is the commutators of the dimensionless conformal vector fields $\boldsymbol{\partial}_a$:

$$[\,\boldsymbol{\partial}_0, \boldsymbol{\partial}_\alpha\,] = \dot{U}_\alpha\,\boldsymbol{\partial}_0 + (q\,\delta_\alpha{}^\beta - \Sigma_\alpha{}^\beta - \epsilon_\alpha{}^\beta{}_\gamma\,R^\gamma)\,\boldsymbol{\partial}_\beta\,, \tag{5}$$

$$[\,\boldsymbol{\partial}_\alpha, \boldsymbol{\partial}_\beta\,] = (2A_{[\alpha}\,\delta_{\beta]}{}^\gamma + \epsilon_{\alpha\beta\delta}\,N^{\delta\gamma})\,\boldsymbol{\partial}_\gamma\,, \tag{6}$$

were $\Sigma_{\alpha\beta}$, $\dot{U}_\alpha$, R_α are the shear, acceleration, and Fermi rotation (which describes how the frame rotates w.r.t. a Fermi propagated frame), respectively, all associated with $\boldsymbol{\partial}_0$ and $\mathbf{G}$; $N^{\alpha\beta}$, A_α are spatial commutator functions that describe the spatial three-curvature of $\mathbf{G}$ (see [11, 26] where the analogous non-normalized objects are described); the object q is the deceleration parameter associated with the physical spacetime and $\mathbf{u}$, but q can also be interpreted geometrically in terms of the expansion Θ of $\boldsymbol{\partial}_0$ according to $q = -\frac{1}{3}\Theta$.

The only variable that carries dimension is H, and hence, for dimensional reasons, the equations associated with H decouple:

$$\boldsymbol{\partial}_0 H = -(1+q)H\,, \qquad \boldsymbol{\partial}_\alpha H = -r_\alpha\,H\,, \tag{7}$$

where the equations also serve to define q and r_α. The dimensionless field equations yield a coupled system of evolution equations and constraints for the variables

$$\mathbf{X} = (E_\alpha{}^i) \oplus \mathbf{S} \oplus \mathbf{M}\,, \qquad \mathbf{S} := (\Sigma_{\alpha\beta}, A_\alpha, N_{\alpha\beta})\,, \tag{8}$$

where $\mathbf{X}$ is the state vector which describes the state space and where $\mathbf{M}$ represents matter variables relevant for the matter source one is interested in, however, from now on, except in the concluding remarks, we will for simplicity restrict ourselves to the vacuum case and hence the state space is described by the state vector $\mathbf{X} = (E_\alpha{}^i) \oplus \mathbf{S}$. In addition to these variables there also exist gauge variables $\mathbf{X}_G$: the R_α Fermi rotation variables are gauge variables associated with the choice of frame; $\dot{U}_\alpha$ is regarded as a gauge variable determined by the choice of $\mathcal{N}$ (or equivalently N). The

---

[7]A special example of such a reference congruence is the synchronous time choice used by BKL.

quantities r_α are generically determined by the Codazzi constraint, however, for certain choices of $\mathcal{N}$ one can derive an evolution equation for r_α, and it may then be advantageous to include r_α in $\mathbf{S}$, see [20]; the deceleration parameter q is determined algebraically by the Raychaudhuri equation, but for certain time choices it is possible and useful to also elevate q to a dependent variable, see [21].

The conformally Hubble-normalized dimensionless system of coupled partial differential equations are obtained from the above Hubble-normalized commutator equations, which define $\mathbf{S}$ and $\mathbf{X}_G$ in terms of derivatives of the metric; the Jacobi identities, which act as integrability conditions for going over to an essentially first order formulation; and Einstein's field equations [26]—the equations can be divided into gauge equations, evolution equations, and constraint equations. The evolution equations for the vacuum case can schematically be written on the form

$$\boldsymbol{\partial}_0 E_\alpha{}^i = F_\alpha{}^\beta E_\beta{}^i \,, \qquad \boldsymbol{\partial}_0 \mathbf{S} = \mathbf{P} \,, \tag{9}$$

where $F_\alpha{}^\beta$ and $\mathbf{P}$ involve $\boldsymbol{\partial}_\alpha, \mathbf{S}, \mathbf{X}_G$ (in the matter case one adds matter equations which yield evolution equations that schematically can be written as $\boldsymbol{\partial}_0 \mathbf{M} = \mathbf{Q}$, but apart from obvious modifications the system takes the same form as in the vacuum case). Note that the Hubble-normalized dimensionless equation system carries the essential dynamical content, since once it is integrated the solution can be inserted into the decoupled dimensional equations for H, which subsequently can be integrated.

To obtain a determined system of evolution equations one needs to specify the spatial frame. There exist many useful spatial frame choices, e.g., the Fermi choice $R_\alpha = 0$; the choice used in [20]; the $SO(3)$ choice used by Benini and Montani as starting point for a 'Misner/Chitre' billiard analysis [28]; and the so-called Iwasawa choice used by Damour, Henneaux, and Nicolai [12]. To establish contact with this latter work we will from now on use the Iwasawa choice, however, it is probably safe to say that no frame choice will cover all the features one may be interested in—each choice has its advantages and disadvantages.

There exists a unique oriented orthonormal spatial frame $\{\omega^\alpha \,|\, \alpha = 1\ldots 3\}$, $\omega^\alpha = e^\alpha_i dx^i$, such that $e^\alpha_i = \sum_\beta \mathcal{D}^\alpha{}_\beta \mathcal{N}^\beta{}_i$,[8] where $\mathcal{D}$ is a diagonal matrix, $\mathcal{D} = \mathrm{diag}\big[\exp(-b^1), \exp(-b^2), \exp(-b^3)\big]$ (to avoid confusion with the Greek frame indices $\alpha, \beta, \ldots$, we use b as the kernel letter for the diagonal degrees of freedom instead of β, which was used in [12]), and $\mathcal{N}^\alpha{}_i$ a unit upper triangular matrix,

$$\left(\mathcal{N}^\alpha{}_i\right) = \begin{pmatrix} 1 & \mathcal{N}^1{}_2 & \mathcal{N}^1{}_3 \\ 0 & 1 & \mathcal{N}^2{}_3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n_1 & n_2 \\ 0 & 1 & n_3 \\ 0 & 0 & 1 \end{pmatrix} \,; \tag{10}$$

$\mathcal{N}^\alpha{}_i$ can also be viewed as representing the Gram-Schmidt orthogonalization of the spatial coordinate coframe $\{dx^i\}$. This choice of frame leads to the so-called Iwasawa decomposition of the spatial metric g_{ij}:

$$g_{ij} = \sum_\alpha \exp(-2b^\alpha) \mathcal{N}^\alpha{}_i \mathcal{N}^\alpha{}_j \,. \tag{11}$$

Using an Iwasawa based conformal Hubble-normalized frame, which we from now on do, leads to $\Sigma_{23} = -R_1, \Sigma_{31} = R_2, \Sigma_{12} = -R_3, N_{33} = 0$ [27]. This makes it natural to introduce the notation $\Sigma_\alpha = \Sigma_{\alpha\alpha}$, while we replace the off-diagonal components of $\Sigma_{\alpha\beta}$ by R_α according to the above equation, i.e.,

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix} = \begin{pmatrix} \Sigma_1 & -R_3 & R_2 \\ -R_3 & \Sigma_2 & -R_1 \\ R_2 & -R_1 & \Sigma_3 \end{pmatrix} \,. \tag{12}$$

---

[8]We here follow the summation conventions in [12]: summation of pairs of coordinate indices $i, j, \ldots$ is understood, whereas sums over the frame indices $\alpha, \beta, \ldots$ are written out explicitly when needed to avoid confusion.

3 The silent boundary

In Section 4 in [20] it is shown that, generically, asymptotic silence is connected with the property that

$$E_\alpha{}^i \to 0\,,$$

while we define the dynamical evolution along a time line to be *asymptotically local* if

$$E_\alpha{}^i \to 0\,, \quad \boldsymbol{\partial}_\alpha(\boldsymbol{S}, r_\beta, \dot{U}_\beta) \to 0\,, \quad (r_\alpha, \dot{U}_\alpha) \to 0\,. \tag{13}$$

towards the singularity. Asymptotic silence and asymptotic local dynamics are closely related concepts, but not necessarily the same. Numerical studies suggest that the dynamical evolution along generic time lines towards generic asymptotically silent singularities is asymptotically local, which corresponds to inhomogeneities being shifted outside shrinking horizons faster than they grow ($E_\alpha{}^i$ goes to zero faster than $\partial_i(\boldsymbol{S}, \log(H), \log(\mathcal{N}))$ may grow), however, numerics also suggest that there may be *special* time lines for which this is not true, which motivates the above distinction [21, 24, 27]. Below, apart from in the concluding remarks, we will discuss the generic case where the evolution is assumed to be asymptotically local.

Consider Einstein's field equations in the conformal Hubble-normalized approach, see Eq. (9), then it follows that

$$E_\alpha{}^i = 0 \tag{14}$$

defines an invariant subset on the boundary of the state space, characterized by the state vector $\boldsymbol{S}$; we refer to this invariant subset as the *silent boundary*. Since $\boldsymbol{\partial}_\alpha = E_\alpha{}^i \partial_i$, Eq. (14) implies that the equations *on* the silent boundary reduce to a system of ordinary differential equations; *the silent boundary can therefore be visualized as an infinite set of copies, parametrized by the spatial coordinates, of a finite dimensional state space*, with $\boldsymbol{S}$ as state vector.

The equations on the silent boundary are identical to the equations for $\boldsymbol{S}$ for spatially self-similar or SH models [29], where the SH case is obtained by setting

$$E_\alpha{}^i = 0 \quad \text{and} \quad \dot{U}_\alpha = 0\,, \, r_\alpha = 0\,. \tag{15}$$

The particular importance of the 'SH silent boundary' in the study of generic spacelike singularities stems from its connection with asymptotically local dynamics; compare Eqs. (13) and (15). Moreover, since the field equations in the conformal Hubble-normalized orthonormal frame approach are regular towards the singularity, we can *extend* the state space to include the silent boundary in our analysis of the dynamical system. This is highly advantageous, since our above reasoning leads to the conjecture that solutions in the full state space generically asymptotically approach and *shadow* the solutions on the SH silent boundary.

The SH silent boundary consists of a number of invariant subsets ('components') that are connected by parts of their boundary only, where the precise structure depends on the choice of frame. Utilizing an Iwasawa frame leads to equations on each component which are the same as those for Bianchi types I–VII, but there is no component associated with Bianchi type VIII or IX. The reason for this is that the Iwasawa frame is incompatible with the symmetry adapted frames of these models.[9] The most general solutions on the SH silent 'Iwasawa' boundary—and these models are as general as the more well known Bianchi type VIII and IX models—are the general Bianchi type $VI_{-1/9}$ solutions [30], and it is thus these models that are of interest as the simplest exact SH example exhibiting generic features in an Iwasawa context.

The asymptotic ODE structure induced by asymptotic local dynamics makes it possible to *reparameterize individual time lines*, for which the spatial coordinates x^i are fixed, so that

$$\boldsymbol{\partial}_0 f = -df/d\tau \tag{16}$$

---

[9]If one uses an Iwasawa frame, Bianchi type VIII and IX solutions appear as inhomogeneous solutions—with their symmetries hidden—in the full interior state space $\boldsymbol{X}$.

holds along a given time line, which allows us to study the asymptotic dynamics along time lines by means of finite dimensional dynamical systems techniques.[10] In the above equation $f(x^0, x^i)$ is any variable occurring in an ODE; $\tau(x^0, x^i) := -\log(\ell/\hat{\ell})$ is a 'local time function' directed towards the singularity; ℓ is related to the the determinant g of the physical spatial metric according to $\ell = g^{1/6}$; $\hat{\ell}$ is an initial value that may vary for different spatial points, i.e., $\hat{\ell}$ is a spatially dependent function, and throughout we will use the convention that hatted objects refer to quantities that are functions of the spatial coordinates alone. To obtain the solution in the chosen time coordinate x^0, one integrates the relation $dx^0 = -\mathcal{N}^{-1}d\tau$ so that

$$x^0 = \hat{x}^0(x^i) - \int_{\tau_0}^{\tau} \mathcal{N}^{-1}(\tau', x^i)d\tau' \,. \tag{17}$$

4   The billiard attractor

If a generic solution approaches the SH part of the silent boundary, then due to the regularity of the equations, it will asymptotically approach the attractor on this subset. Our experience with SH cosmology suggests that this attractor is somehow connected with the Bianchi type I and II subsets on the silent boundary, but that it also depends on the choice of frame.

Dynamical systems investigations usually begin by looking for fixed points (equilibrium points) and then one subsequently linearizes the equations at these points to obtain a local picture of dynamical features. In the Iwasawa gauge there exists a representation of the (vacuum) Bianchi type I, i.e. Kasner, solutions on the silent boundary as a one-parameter set of fixed points, which play a key role for the asymptotic dynamics; this set of fixed points is referred to as the *Kasner circle* $\mathrm{K}^{\bigcirc}$, and is determined by

$$1 - \Sigma^2 = A_\alpha = N_{\alpha\beta} = R_\alpha = 0 \,, \tag{18}$$

where $\Sigma^2 = \frac{1}{6}\Sigma_{\alpha\beta}\Sigma^{\alpha\beta}$. It follows from Eqs. (12) and (18) that $\Sigma_{\alpha\beta} = \mathrm{diag}[\hat{\Sigma}_1, \hat{\Sigma}_2, \hat{\Sigma}_3]$, however, one usually represents the Kasner circle by the generalized Kasner exponents p_α:

$$\Sigma_{\alpha\beta} = \mathrm{diag}[\hat{\Sigma}_1, \hat{\Sigma}_2, \hat{\Sigma}_3] = \mathrm{diag}[3p_1 - 1, 3p_2 - 1, 3p_3 - 1] \,, \tag{19}$$

where we omit the hats on top of p_α in order to agree with standard notation. Since $\mathrm{tr}\,\Sigma_{\alpha\beta} = 0$ and $\Sigma^2 = 1$, the Kasner exponents satisfy the Kasner relations

$$p_1 + p_2 + p_3 = 1 \,, \qquad p_1^2 + p_2^2 + p_3^2 = 1 \,, \tag{20}$$

which imply that $-\frac{1}{3} \le p_\alpha \le 1$ (or equivalently, $-2 \le \Sigma_\alpha \le 2$).

The Kasner circle $\mathrm{K}^{\bigcirc}$ can be divided into six equivalent sectors, identifiable by means of permutations of the spatial axes. Each sector can be characterized by an ordered sequence of Kasner exponents; hence e.g. sector (312) is defined as the part of $\mathrm{K}^{\bigcirc}$ where $p_3 < p_1 < p_2$, see Fig. 4. There are six special points on $\mathrm{K}^{\bigcirc}$ that are associated with solutions with additional symmetries: the points Q_α correspond to the three equivalent non-flat plane symmetric solutions, while the points T_α correspond to the Taub representation of Minkowski spacetime, see Fig. 4.[11]

Linearization of the dimensionless field equations at $\mathrm{K}^{\bigcirc}$ yields a set of ODEs that tells us that $E_\alpha{}^i$, $N_{\alpha\beta}$ ($\alpha \ne \beta$) and A_α belong to the stable subspace of each fixed point of $\mathrm{K}^{\bigcirc}$ (except at the Taub points; e.g. $\partial_\tau E_\alpha{}^i = -3(1 - p_\alpha)E_\alpha{}^i$), while R_α and $N_1 = N_{11}, N_2 = N_{22}$ are stable or unstable depending on the sector of $\mathrm{K}^{\bigcirc}$ the point (p_1, p_2, p_3) lies in. Finally, the variables $\Sigma_\alpha = \Sigma_{\alpha\alpha}$ belong to the center subspace, i.e., they are constant to first order. The analysis of the stability of the Kasner circle $\mathrm{K}^{\bigcirc}$ is summarized in Fig. 4, where the unstable variables are given for each sector of $\mathrm{K}^{\bigcirc}$.

---

[10]This is analogous to the dynamical systems treatment of the so-called silent models, see [31, 32] and [11] [Chapter 13]; in the present context it is irrelevant that the only nontrivial spatially inhomogeneous non-rotating silent solutions without a cosmological constant turned out to be the Szekeres dust models, see [33, 34], and references therein.

[11]For Q_1 we have $(\hat{\Sigma}_1, \hat{\Sigma}_2, \hat{\Sigma}_3) = (-2, 1, 1)$ or $(p_1, p_2, p_3) = (-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$; T_1 is characterized by $(\hat{\Sigma}_1, \hat{\Sigma}_2, \hat{\Sigma}_3) = (2, -1, -1)$ or $(p_1, p_2, p_3) = (1, 0, 0)$; Q_2, Q_3, T_2, T_3 are obtained by cyclic permutations).
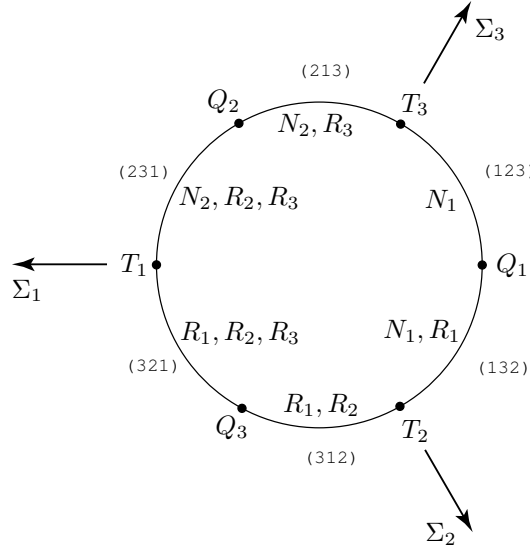
Figure 1: The Kasner circle of fixed points—sectors and unstable variables.

The linear analysis suggests that the stable variables $E_\alpha{}^i$ and $\boldsymbol{S}_{\text{stable}} = (N_{\alpha\beta}, A_\gamma)$ $(\alpha \neq \beta)$ decay rapidly, and since a generic solution is expected to spend an increasing amount of time in a small neighborhood of the 'non-flat' part of K°, we are led to the conjecture that a generic solution must approach the invariant *oscillatory subset* $\mathcal{O}$, determined by $\boldsymbol{S}_{\text{stable}} = 0$ on the SH silent boundary. This motivates decomposing the state vector $\boldsymbol{S} = (\Sigma_{\alpha\beta}, A_\alpha, N_{\alpha\beta})$ of the SH silent boundary into a 'stable' and an 'oscillatory' part:

$$\boldsymbol{S} = \boldsymbol{S}_{\text{stable}} \oplus \boldsymbol{S}_{\text{osc}} \qquad \text{where} \qquad \boldsymbol{S}_{\text{osc}} = (\Sigma_\alpha, R_\alpha, N_1, N_2) \,. \tag{21}$$

The oscillatory subset $\mathcal{O}$, characterized by the state vector $\boldsymbol{S}_{\text{osc}}$, consists of components describing various representations of Bianchi type I, II, VI$_0$, VII$_0$ solutions. All the orbits, i.e. solutions, on these components are heteroclinic orbits, i.e. they originate and end at fixed points.

Since $E_\alpha{}^i \to 0, \boldsymbol{S}_{\text{stable}} \to 0$, it is the asymptotic behavior of the remaining oscillatory variables $\boldsymbol{S}_{\text{osc}}$ that represents the nontrivial asymptotic dynamics of a generic solution $\boldsymbol{X}(\tau)$, at a generic spatial point x^i, as $\tau \to \infty$. Moreover, due to the heteroclinic orbit structure on $\mathcal{O}$, we expect $\boldsymbol{S}_{\text{osc}}(\tau)$ to be increasingly accurately described by a *partition based on a sequence of segments, where each segment is associated with a heteroclinic orbit on $\mathcal{O}$ and where two subsequent segments of $\boldsymbol{S}_{\text{osc}}(\tau)$ are joined at a point of* K°, yielding an oscillatory behavior. We refer to this description of the orbit $\boldsymbol{S}_{\text{osc}}(\tau)$ in the asymptotic regime as an *asymptotic sequence of $\mathcal{O}$-orbits $\mathcal{AS}_\mathcal{O}$*, or for brevity, as an *asymptotic sequence* (for an example of a part of an asymptotic sequence, see Fig. 3(b)). Note that the concept of asymptotic sequences $\mathcal{AS}_\mathcal{O}$ generalizes, connects, and gives precise meaning to BKL's 'piecewise approximations'.

The heteroclinic orbits on $\mathcal{O}$ form a 'heteroclinic orbit puzzle' that governs asymptotically local dynamics, but an analysis shows that it is only those orbits that connect points on K° that are of relevance for asymptotic generic dynamics [27]; these heteroclinic orbits can be regarded as representing *transitions* between different points on K°.

Further analysis shows that generic dynamics is associated with a subset of $\mathcal{O}$ —the *billiard attractor* subset $\mathcal{O}_{\mathcal{BA}}$ [27], defined as

$$\mathcal{O}_{\mathcal{BA}} = \mathrm{K}^\circ \cup \mathcal{B}_{N_1} \cup \mathcal{B}_{R_1} \cup \mathcal{B}_{R_3} \,, \tag{22}$$

where $\mathcal{B}_{N_1}, \mathcal{B}_{R_1}, \mathcal{B}_{R_3}$ are subsets on $\mathcal{O}$ characterized by $N_2 = R_1 = R_2 = R_3 = 0$ (the silent Bianchi type II subset associated with N_1 and a Fermi frame), $N_1 = N_2 = R_2 = R_3 = 0$ (the silent Kasner subset in a frame that rotates w.r.t. a Fermi propagated frame in the 2-3-plane), $N_1 = N_2 = R_1 = R_2 = 0$ (the silent Kasner subset in a frame that rotates w.r.t. a Fermi

propagated frame in the 1-2-plane), respectively; transitions, i.e. heteroclinic orbits that connect points on K°, associated with these subsets, denoted by $\mathcal{T}_{N_1}$ (called single curvature transitions), $\mathcal{T}_{R_1}$, $\mathcal{T}_{R_3}$ (called single frame transitions), are depicted as projections onto Σ_α-space in Fig. 2.
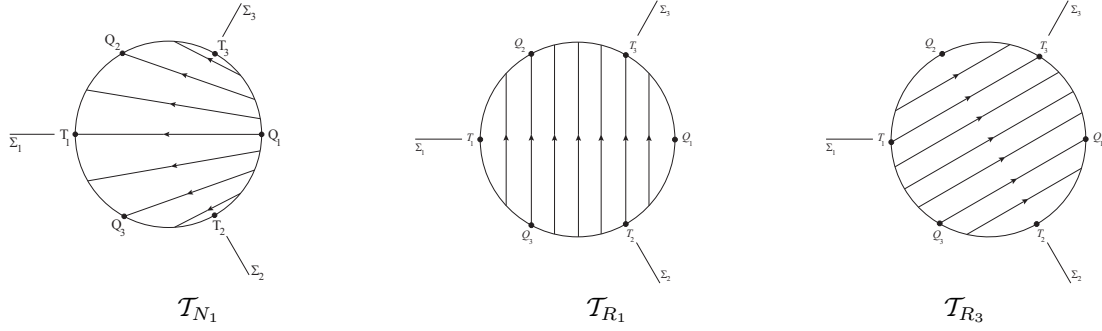

Figure 2: Projections onto diagonal Σ-space of the single transitions $\mathcal{T}_{N_1}$, $\mathcal{T}_{R_1}$, $\mathcal{T}_{R_3}$ on the billiard attractor for the Iwasawa frame.

An *attractor sequence* $\mathcal{A}_\mathcal{T}$ is a sequence of transitions on the billiard attractor $\mathcal{O}_{\mathcal{BA}}$, i.e an infinite concatenation of single transitions $\mathcal{T}_{N_1}$, $\mathcal{T}_{R_1}$, and $\mathcal{T}_{R_3}$; to get an intuitive picture of attractor sequences we refer to Fig. 2 and give an example of part of an attractor sequence in Fig. 3(*b*).

The above leads to the formulation of the *dynamical systems billiard conjecture* [27]:

Conjecture. *The asymptotic dynamics of a generic time line of a solution of Einstein's vacuum equations (expressed in an Iwasawa frame) that exhibits a generic spacelike singularity is characterized as follows:*

(i) *It is asymptotically silent and local.*

(ii) *In the asymptotic limit the essential dynamics is represented by an attractor sequence $\mathcal{A}_\mathcal{T}$ on the billiard attractor $\mathcal{O}_{\mathcal{BA}}$.*

Let us comment on this, and in particular, on some of the features and ingredients in the *derivation* in [27] that led to the above conjecture:

(i) The suppression of a priori possible generic behavior can be divided into two kinds: a) 'dynamical' suppression, b) 'stochastical' suppression.

(ii) The derivation of 'suppression estimates' is based on utilizing the fact that the asymptotic dynamical evolution of a generic spatial point x^i of a generic solution $\boldsymbol{X}(\tau)$ shadows a sequence of orbits/transitions on $\mathcal{O}$ with an increasing degree of accuracy. This leads to the concept of asymptotic sequences, which yield a simple asymptotic dynamical description in the context of the full extended state space picture.

(iii) Although an orbit is increasingly accurately approximated by a sequence of heteroclinic orbit segments, decreasing errors nevertheless occur, and from a stochastic point of view this is essential: they lead to a 'randomization of heteroclinic orbits,' and this is what makes a stochastic analysis admissible.

(iv) The stochastic analysis of sequences of transitions relies on a generalization of BKL's concept of so-called eras; sequences are partitioned into *small curvature phases* (series of heteroclinic orbits close to the Taub points where the curvature is small) and *large curvature phases* (series of heteroclinic orbits far from the Taub points where the curvature is large), where, stochastically, small curvature phases turn out to dominate over large curvature phases.

(v) The dynamical and stochastical analysis leads to an estimation of decay rates which shows that, in the context of a hierarchy of subsets, the asymptotic dynamical evolution is restricted to subsets of subsets, to boundaries of boundaries, descending from the full state space via the SH silent boundary and the oscillatory subset down to the billiard attractor.

(vi) Stochastic considerations are crucial in the derivation of the billiard attractor. The possibility therefore exists that there are solutions with different asymptotic behavior, but such solutions are expected to form a set of measure zero in the space of all solutions.

(vii) The attractor is a local attractor in the full state space—there exist open sets of solutions not approaching it (e.g., an open set of solutions that are almost flat everywhere).

The derivation of the billiard conjecture in [27] is not mathematically rigorous, but depends on arguments that are heuristic—despite their being convincing. However, it is reasonable to expect that if one attempts to obtain proofs, then several concepts and methods introduced in [27] will play a prominent role. Let us give an example: The underlying structures that allowed one to obtain mathematical proofs about the attractor of the diagonal Bianchi type IX models in a Fermi propagated frame [17, 18] are specific for these models and are not available in other cases; since it is such cases that are relevant for the present general scenario, the diagonal Bianchi type IX models are misleading. The analysis in [27] suggests that one has to know the (asymptotic) history of a solution to unravel its asymptotic features; this causes a dilemma since this requires that one finds the solution, which seems unlikely. However, randomization makes it possible to stochastically examine the cumulative effects of small and large curvature phases and this allows one to estimate decay rates and give a description of what is going to happen generically—statistical analysis is one example of an ingredient that is likely to play a role in future proofs.

5   'Duality' of Hamiltonian and dynamical systems billiards

In the Hamiltonian approach by Damour, Henneaux, and Nicolai [12, 13] it is shown that the essential generic asymptotic dynamics can be described by a Hamiltonian cosmological billiard. In this picture the evolution of a spatial point of a (generalized) Fermi propagated Kasner solution appears as a geodesic in hyperbolic space, but this space is bounded by sharp walls, and this leads to bounces, see Fig. 3(a); the walls and bounces are of two kinds: (i) frame/centrifugal/symmetry walls and associated bounces; (ii) curvature walls and associated bounces. Bounces of the former type merely result in axis permutations of a Kasner solution, while bounces of the latter type lead to a change in the Kasner state generated by a Bianchi type II solution.

In the dynamical systems formulation free motion in the Hamiltonian picture corresponds to the points on the Kasner circle $\mathrm{K}^{\bigcirc}$; centrifugal bounces in the Hamiltonian approach translate to single frame transitions $\mathcal{T}_{R_1}$ and $\mathcal{T}_{R_3}$, while curvature bounces are connected with single curvature transitions $\mathcal{T}_{N_1}$. Here we observe 'bounces' at the fixed points on the Kasner circle $\mathrm{K}^{\bigcirc}$, which act as a 'wall', and between bounces we have motion along straight lines in (Euclidean) Σ_α-space (transitions), see Fig. 3(b).

Thus the Hamiltonian 'motion–bounce–motion–bounce' is replaced with 'bounce–motion–bounce–motion' in the dynamical systems billiards. In the Hamiltonian approach the essential asymptotic dynamical evolution is described by 'configuration space' variables (related to diagonalized spatial metric variables) while in the dynamical systems approach it is described in terms of 'momentum space' variables (Σ_α are related to time derivatives of the diagonalized spatial metric, and thus to the associated momenta). The two approaches thus give *complementary 'dual' representations* of the generic asymptotic dynamics; to compare the two pictures, see Fig. 3. In addition the two approaches mutually support each other and strengthen the credibility of the generic picture; moreover, the combined use of both may very well be of relevance for future developments, see [27] for further discussion.
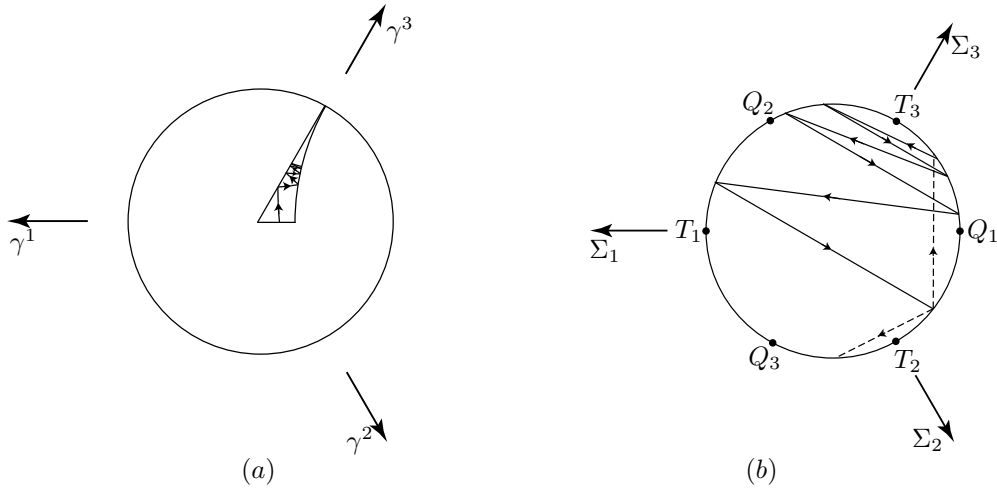
Figure 3: Fig. (*a*) shows part of an orbit at a spatial point in terms of free Kasner (Fermi frame) motion and frame and curvature bounces. This represents a 'configuration space' projection of the asymptotic dynamics. The disc here represents hyperbolic space. (The γ^α variables are 'orthogonal' angular variables associated with a projection in b^α-space, see [12, 27].) Fig. (*b*) shows part of an orbit in terms of single frame and curvature transitions, i.e., it shows a part of an attractor sequence. Note that the solution does not quite return to any of the Kasner points it has 'visited' before. This description represents a 'momentum space' projection of the asymptotic dynamics. The circle here is the Kasner circle K°. The dashed lines correspond to the two possible single transitions that are possible at this stage; which one is realized depends on initial data. This corresponds to the situation that free motion in a given direction in Fig. (*a*) may lead either to hitting the wall associated with R_1 (the short wall) or with N_1 (the curved wall).

6 Concluding remarks

In this paper we have presented the conformal Hubble-normalized dynamical systems approach, and outlined the derivation of the billiard attractor and established its 'dual' correspondence with the Hamiltonian cosmological billiard of Damour et al. [12], but it is also of interest to make a comparison with BKL: the generalized Kasner solution is obtained by inserting the values of **S** at the Kasner circle K° into the equations for $E_\alpha{}^i$ and H, and thus this 'solution' is obtained as the lowest order perturbation of K° into the physical state space ($E_\alpha{}^i \neq 0$); similarly the generalized Bianchi type II solutions correspond to single curvature transitions $\mathcal{T}_{N_1}$—the ad hoc starting point of BKL is thereby derived. Furthermore, as discussed, asymptotic (attractor) sequences connect and yield precise meaning to BKL's 'piecewise approximations,' and thus the results of BKL's analysis are rigorously formulated as special structures in the full extended dynamical systems picture.

Here we have been concerned with the vacuum case. However, the methods employed in [27] are equally applicable to many other theories, see e.g. [12, 13], and to sources as well. Regarding possible sources this naturally leads to a study of the influence of matter on generic spacelike singularities, i.e. structural stability of generic singularities, and this in turn leads to a variety of issues, e.g., can the following be made more substantial? Generic singularities seem to bring out the essential properties of matter. There are indications that it is natural to base a classification of the influence of matter on generic singularity structure on (i) energy conditions and (ii) whether the effective propagation speed is less than that of light or not. There are also indications that suggest a subclassification based on how 'matter matters' in the light speed case, e.g. massless scalar fields and electromagnetic fields influence the generic spacelike singularity in different ways; does this motivate a subclassification based on spin? If we assume the usual energy conditions and consider the case of less than light speed propagation, then 'matter does not seem to matter'

for generic asymptotic dynamics towards a generic singularity—solutions are said to be 'vacuum dominated', or more precisely, the Hubble-normalized energy-momentum tensor tends to zero so that the geometry is asymptotically described by vacuum solutions. However, this does not necessarily mean that e.g. the Hubble-normalized rotation of a perfect fluid tends to zero, which suggests a subclassification based on features that are (if any), or are not, affected by matter. Another issue is how and if matter influences the connection between generic spacelike singularities and weak null singularities in asymptotically flat spacetimes (there are indications that e.g. Vlasov matter behaves differently than perfect fluids [35]); for special examples, see [25].

In this paper we have assumed that the dynamical evolution along generic time lines is associated with asymptotic silence and asymptotic local dynamics, which is supported by the results in [27], but this does not mean that there could not exist interesting phenomena associated with *special* time lines, indeed, we believe that a set of time lines of measure zero exhibits spike formation and recurring 'spike transitions' [24], associated with non-local dynamics and the formation of large spatial gradients—but asymptotic silence still holds. In this context it is important to note that even though spatial partial derivatives grow without bounds, the Hubble-normalized spatial frame derivatives and the dimensionless state space variables are still bounded, i.e., asymptotic regularization still holds! Moreover, spike transitions seem to be connected with asymptotically local dynamics and are governed by variations of the Kasner map, hinting at further more deeply hidden structures. Asymptotic spike formation is associated with the fact that the unstable variables N_1, R_1, R_3 go through a zero at a spatial surface, but the underlying mechanisms of spike formation and possible spike annihilation are not completely understood [36]: Where and why do spikes form? Do there exist spikes that persist and lead to infinite sequences of spike transitions (infinite recurring spike transitions)? Do spikes annihilate spikes, and if so how: through 'spike interference'? Is spike formation more common than spike annihilation so that 'spike cascading' occurs? How many spikes form: do they form a dense set? Do spikes—'gravitational defect-like' surfaces from the very early universe—leave possible observational imprints in e.g. the CMB? It clearly is of interest to answer these questions, moreover, a clarification of some of these issues is of considerable interest in the context of eventual generic singularity proofs.

Finally, one may ask why one should study generic singularities at all in a classical GR context? Firstly, there exists a regime between the Planck and GUT eras in the very early universe where GR is expected to hold and where the approach towards the initial singularity presumably is described by the dynamics towards a generic singularity (recall that one of the points of inflation is to 'erase the effects of initial data,' and that before this erasure a singularity is presumably generic according to this line of reasoning). Secondly, black hole formation is connected with initial data that reflects the complexities of the real universe; one would hence also in this case expect generic spacelike singularities to play a role before one enters the Planck regime. Thirdly, the formation of generic singularities is associated with considerable structure (for the exploitation of some of these structures in the context of quantization of special models, see e.g. [37, 38]), even in the case of spike formation: Can this structure be used to asymptotically quantize gravity where it needs to be quantized, namely in the ultra strong gravitational field in the neighborhood of a generic spacelike singularity?

Acknowledgements

References

[1] E.M. Lifshitz and I.M. Khalatnikov. *Adv. Phys.* **12**, 185 (1963).

[2] V.A. Belinskiĭ, I.M. Khalatnikov, and E. M. Lifshitz. *Adv. Phys.* **19**, 525 (1970).

[3] V.A. Belinskiĭ, I.M. Khalatnikov, and E.M. Lifshitz. *Adv. Phys.* **31**, 639 (1982).

[4] I.M. Khalatnikov, E.M. Lifshitz, K.M. Khanin, L.N. Shur, and Ya.G. Sinai. *J. Stat. Phys.* **38**, 97 (1985).

[5] J.D. Barrow. *Phys. Rep.* **85**, 1 (1982).

[6] D.F. Chernoff and J.D. Barrow. *Phys. Rev. Lett.* **50**, 134 (1983).

[7] C.W. Misner. *Phys. Rev.* **186** 1319 (1969).

[8] C.W. Misner, K.S. Thorne, and J.A. Wheeler. *Gravitation.* (W.H. Freeman and Company, San Francisco, 1973).

[9] D.M. Chitré. *Ph.D. Thesis.* University of Maryland, (1972).

[10] R.T. Jantzen. in *Proc. Int. Sch. Phys. "E. Fermi" Course LXXXVI on "Gamov Cosmology"*, R. Ruffini, F. Melchiorri, Eds. (North Holland, Amsterdam, 1987) and in *Cosmology of the Early Universe*, R. Ruffini, L.Z. Fang, Eds. (World Scientific, Singapore, 1984). arXiv:gr-qc/0102035.

[11] J. Wainwright and G.F.R. Ellis. *Dynamical systems in cosmology.* (Cambridge University Press, Cambridge, 1997).

[12] T. Damour, M. Henneaux, and H. Nicolai. *Class. Quantum Grav.* **20** R145 (2003).

[13] T. Damour and H. Nicolai. *Class. Quantum Grav.* **22** 2849 (2005).

[14] B.K. Berger, D. Garfinkle, J. Isenberg, V. Moncrief, and M. Weaver. *Mod. Phys. Lett.* **A13** 1565 (1998).

[15] B.K. Berger. *Living Reviews in Relativity.* **5** 6 (2002).

[16] L. Andersson and A.D. Rendall *Commun. Math. Phys.* **218** 479 (2001).

[17] H. Ringström. *Class. Quantum Grav.* **17** 713 (2000).

[18] H. Ringström. *Annales Henri Poincaré* **2** 405 (2001).

[19] A.A. Coley. *Dynamical systems and cosmology.* (Kluwer Academic Publishers 2003).

[20] C. Uggla, H. van Elst, J. Wainwright, and G.F.R. Ellis. *Phys. Rev. D* **68** : 103502 (2003).

[21] D. Garfinkle. *Phys. Rev. Lett.* **93** 161101 (2004).

[22] W.C. Lim. *The Dynamics of Inhomogeneous Cosmologies* Ph.D. thesis, (University of Waterloo 2004); arXiv:gr-qc/0410126.

[23] W.C. Lim, H. van Elst, C. Uggla, and J. Wainwright. *Phys. Rev. D* **69** : 103507 (2004).

[24] L. Andersson, H. van Elst, W.C. Lim, and C. Uggla. *Phys. Rev. Lett.* **94** 051101 (2005).

[25] W.C. Lim, C. Uggla, and J. Wainwright. *Class. Quantum Grav.* **23** 2607 (2006).

[26] N. Röhr and C. Uggla. *Class. Quantum Grav.* **22** 3775 (2005).

[27] J.M. Heinzle, C. Uggla, and N. Röhr. The cosmological billiard attractor. arXiv:gr-qc/0702141.

[28] R. Benini and G. Montani. *Phys. Rev. D.* **70** 103527-1 (2004).

[29] L. Andersson, H. van Elst, and C. Uggla. *Class. Quantum Grav.* **21** S29 (2004).

[30] C.G. Hewitt, J.T. Horwood, and J. Wainwright. *Class. Quantum Grav.* **20** 1743 (2003).

[31] S. Matarrese, O. Pantano, and D. Saez. *Phys. Rev. Lett.* **72** 320 (1994)

[32] M. Bruni, S. Matarrese, and O. Pantano. *Astrophys. J.* **445** 958 (1995)

[33] H. van Elst, C. Uggla, W. Lesame, G.F.R. Ellis, and R. Maartens. *Class. Quantum Grav.* **14** 1151 (1997).

[34] L. Wylleman and N. Van den Bergh. arXiv:gr-qc/0604025.

[35] J.M. Heinzle and C. Uggla. *Class. Quantum Grav.* **23** 3463 (2006).

[36] W.C. Lim. Private communication.

[37] A. Ashtekar, R.S. Tate, and C. Uggla. *Int. J. Mod. Phys.* **D2** 15 (1993).

[38] A. Ashtekar, R.S. Tate, and C. Uggla. Minisuperspaces: Symmetries and Quantization. In: Misner Festschrift, Edited by B.L. Hu et al., (Cambridge University Press 1993); arXiv:gr-qc/9302026.